11/21 Solutions

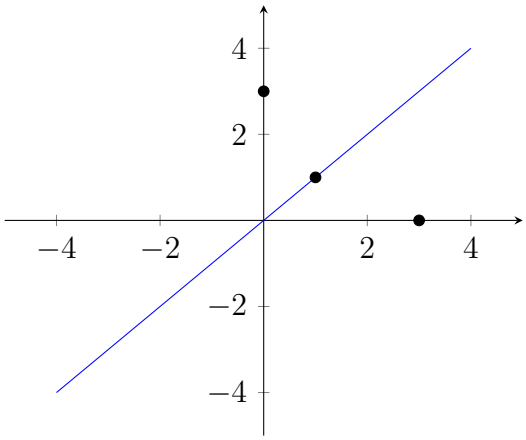
1. Consider the vectors $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \in \mathbb{R}^2$. Let $A = \begin{bmatrix} 3 & 0 \\ 1 & 1 \\ 0 & 3 \end{bmatrix}$.

   Consider the following SVD of A

$$\begin{bmatrix} 3 & 0 \\ 1 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 3\sqrt{22}/22 & -\sqrt{2}/2 & \sqrt{11}/11 \\ \sqrt{22}/11 & 0 & -3\sqrt{11}/11 \\ 3\sqrt{22}/22 & \sqrt{2}/2 & \sqrt{11}/11 \end{bmatrix} \begin{bmatrix} \sqrt{11} & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}^T$$

   (a) Find and sketch the best fit 1−dimensional subspace of $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$. What quantity does this line minimize?

   The best fit 1−dimensional subspace of the data points is given by the span of the first singular vector; that is $\text{Span}\{\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}\}$. This is represented by the line $y = x$ in $\mathbb{R}^2$, sketched below.



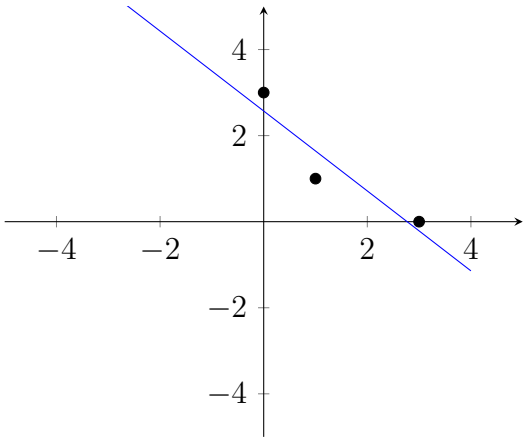
   This line minimizes the sum of the squares of the perpendicular distances between the subspace and the data points.

   (b) Find and sketch the least squares line of best fit of $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$. What quantity does this line minimize?

   A line of the form $y = mx + b$ gives rise to the following system (with no solution):

$$\begin{aligned} 3m + b &= 0 \\ m + b &= 1 \\ b &= 3 \end{aligned}$$

giving us the matrix equation $\begin{bmatrix} 3 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$. To find a least squares solution, we must solve $\begin{bmatrix} 3 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 3 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ or $\begin{bmatrix} 10 & 4 \\ 4 & 3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$. This has solution $\mathbf{x} = \begin{bmatrix} -\frac{13}{14} \\ \frac{18}{7} \end{bmatrix}$. Hence $y = -\frac{13}{14}x + \frac{18}{7}$ is the least squares line of best fit, drawn below.



This line minimizes the sum of the squares of the vertical distances between the line and the data points.

(c) Which of the two lines gives a better approximation of the data? What feature of the line found in part (b) precludes it from being the best fit 1−dimensional subspace?

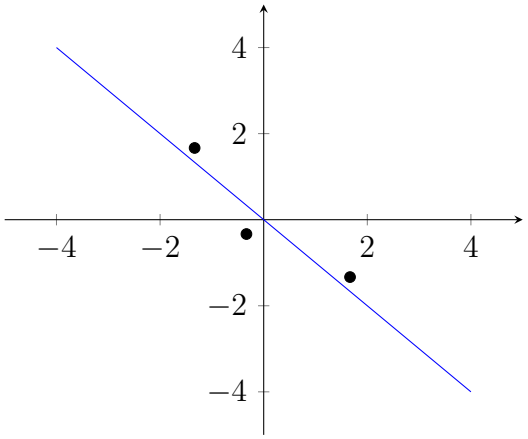
The line in part (b) gives a better approximation. Since the line in part (b) does not pass through the origin, it is not a subspace of $\mathbb{R}^2$, so it cannot be the 1-dimensional subspace of best fit.

(d) **Looking Ahead** If we *center* the data about the origin, then the best fit 1−dimensional subspace will provide a more accurate line of best fit. This is what is done in *principle component analysis* (which we will discuss next week). Let's do this in this example. Let A' be the matrix obtained by subtracting each entry of the ith column of A by the average of the entries in the ith column of A. This yields the matrix

$$A' = \begin{bmatrix} 5/3 & -4/3 \\ -1/3 & -1/3 \\ -4/3 & 5/3 \end{bmatrix}$$

Plot the related data points in $\mathbb{R}^2$ and compare it to the old data points; notice that they look like the old data points shifted towards the origin. Next, use Julia to find an SVD of A' and use it to find the best fit 1-dimensional subspace; sketch it.

Using Julia, we see that the first singular vector is $\begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$. Hence the best fit subspace is Span $\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$, which is represented by the line $y = -x$, shown below.



2. Best fit subspaces are not necessarily unique. Consider the following data points in $\mathbb{R}^3$.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

(a) Find an SVD of the matrix A whose rows are $\mathbf{x}_i^T$ (by hand).

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. Then $A^TA = I_3$ is the 3×3 identity matrix. It has a single eigenvalue 1, with multiplicity 1. Since it is diagonalizable, $E_1(A^TA)$ has orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Hence $V = I_3$ and $\Sigma = A$. Now the first three columns of U are given by

$$\frac{A\mathbf{e}_1}{1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \frac{A\mathbf{e}_2}{1} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \text{ and } \frac{A\mathbf{e}_3}{1} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

The final column of U is a unit basis vector of $N(A^T)$, which can be shown to be

$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ (check). Hence $U = I_4$, the 4×4 identity. Hence $A = I_4 A I_3$ is an SVD.

(b) Determine (at least) two best fit 2-dimensional subspaces.

Since all three columns of V are eigenvectors corresponding to 1, any two can be used as a basis for a $2-$dimensional subspace of best fit. Two such examples are

$$W_1 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

$$W_2 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

(c) Visit `https://www.geogebra.org/m/qsfprdt3` to see two such $2-$dimensional best fit subspaces. Determine $\sum ||e_i||^2$ (sum of squares of the error vectors) for each plane (geometrically) and verify that they give the same value.

For each of the two planes $\sum ||e_i||^2 = 1$. Note that in each case, the plane P contains exactly three of the vectors and the fourth vector is perpendicular to P. Hence $\sum ||e_i||^2 = 1$ is simply the length squared of the fourth vector, which has length 1.